

An Update on CORBA Performance for HPEC Algorithms

Bill Beckwith

Objective Interface Systems, Inc.

Email: bill.beckwith@ois.com

CORBA technology today surrounds HPEC-oriented subsystems. In recent years CORBA is getting used inside those subsystems, but mostly to facilitate communication with nodes outside the subsystem. It is now possible to implement CORBA ORBs that have the performance characteristics required by HPEC applications. This talk will describe the effort to modify the OMG CORBA specification to accommodate HPEC requirements, one HPEC ORB implementation, and preliminary measured performance data.

CORBA provides a standards-based middleware architecture for building flexible distributed systems. The time-to-market and engineering life-cycle benefits of using CORBA in enterprise, server, and desktop systems are well documented. At the very least, embedding CORBA in very high-performance and parallel computing environments offers seamless connectivity to external environments such as Java virtual machines, web-integrated application agents, etc. Beyond the basic value, potential exists for building time-critical, data-intensive applications with the more flexible CORBA programming paradigm where communication is handled by a highly specialized Object Request Broker (ORB).

The determinate of this potential is the performance impact of using ORB technology on these time-critical, data-intensive applications. With the historical performance of desktop ORBs we often expect the CORBA GIOP protocol, and thus ORBs by nature, must add significant overhead to simple communications.

Thus, the wide-spread acceptance of CORBA use in very high-performance and parallel computing environments is predicated on the existence of specialized ORB technologies that can achieve nominal latency and optimal throughput consistent with or better than other parallel computing middleware technologies.

A useful timeliness measure of any ORB technology is the degree to which the use of the ORB is transparent with respect to the application performance. A parallel-computing ORB that provides a highly efficient use of the latest interconnect hardware makes the use of the ORB temporally transparent to the application.

The performance threshold that makes an ORB technology temporally transparent to applications depends on the nature of the application. An ORB that adds hundreds of microseconds to a small message transmission would not offer temporal transparency for most applications, but might offer temporal transparency for systems with less stringent requirements. If, however, an ORB technology can achieve performance by directly accessing hardware features of a high-speed interconnect that surpasses the performance of typical direct usage then the ORB become practical for all parallel-computing application uses.

Report Documentation Page			Form Approved OMB No. 0704-0188	
<p>Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p>				
1. REPORT DATE 20 AUG 2004	2. REPORT TYPE N/A	3. DATES COVERED -		
4. TITLE AND SUBTITLE An Update on CORBA Performance for HPEC Algorithms			5a. CONTRACT NUMBER	
			5b. GRANT NUMBER	
			5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)			5d. PROJECT NUMBER	
			5e. TASK NUMBER	
			5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Objective Interface Systems, Inc.			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSOR/MONITOR'S ACRONYM(S)	
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release, distribution unlimited				
13. SUPPLEMENTARY NOTES See also ADM001694, HPEC-6-Vol 1 ESC-TR-2003-081; High Performance Embedded Computing (HPEC) Workshop (7th)., The original document contains color images.				
14. ABSTRACT				
15. SUBJECT TERMS				
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 17
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified		

There are two elements of the performance of a communication technology that are useful to this discussion: latency and throughput. Latency is typically measured in fractions of seconds and represents the time delay from when the sending begins to initiate a data transfer to the point at which the recipient starts to receive that data. Latency is easily benchmarked by measuring the end-to-end time to deliver small messages. Throughput is typically measured in bytes per second and represents a measure of the utilization of the underlying hardware's communication bandwidth.

A performant ORB technology would ideally:

- (1) add little latency to low-latency, high-speed interconnect technologies for small messages, and
- (2) add little or no overhead per byte transferred to high-speed interconnect technology for large data transmissions (i.e. minimal bandwidth reduction).

These two metrics are nicely represented on a X-Y graph where the X axis is data transmission size and the Y axis is total transmission time. A useful latency metric is the Y intercept of the line. A useful metric for throughput is the slope of the line.

The coordinated engineering of a highly efficient ORB implementation and high-speed interconnect hardware can offer application architects performance superior to the alternative of custom-designing the application. However, this superiority is only possible if the ORB technology is purpose-built and very closely integrated with the high-speed interconnect hardware.

Additionally, since such an application is using standard conformant APIs the application is portable to environments other than the performant high-speed interconnect and future generations of high-speed interconnects.

Latency and throughput provide only a piece of the performance puzzle. The correctness of many real-time applications depends on the predictability of the latency and throughput. A full discussion of optimality criteria is beyond the scope of this discussion but is important to application engineers building these systems and as a design constraint on the underlying ORB technology.



An Update on CORBA Performance for HPEC Algorithms

Bill Beckwith
Objective Interface Systems, Inc.

bill.beckwith@ois.com

<http://www.ois.com>

September 2nd, 2003



Elements of Performance

- ◆ **Simplified (but accurate) execution model:**
 - Latency
 - End-to-end time to transfer one byte
 - Per Byte
 - Extra end-to-end time to transfer each additional byte
 - Total time
 - Latency + Per Byte * Bytes
- ◆ **Copies add to Per Byte time**
 - HPEC hardware transfer rates are competitive with local memcpy times (approx. one byte per clock cycle)
 - Result is that any copies kill throughput (but you knew that)



First Benchmarks: Zero Copy Affect on Windows

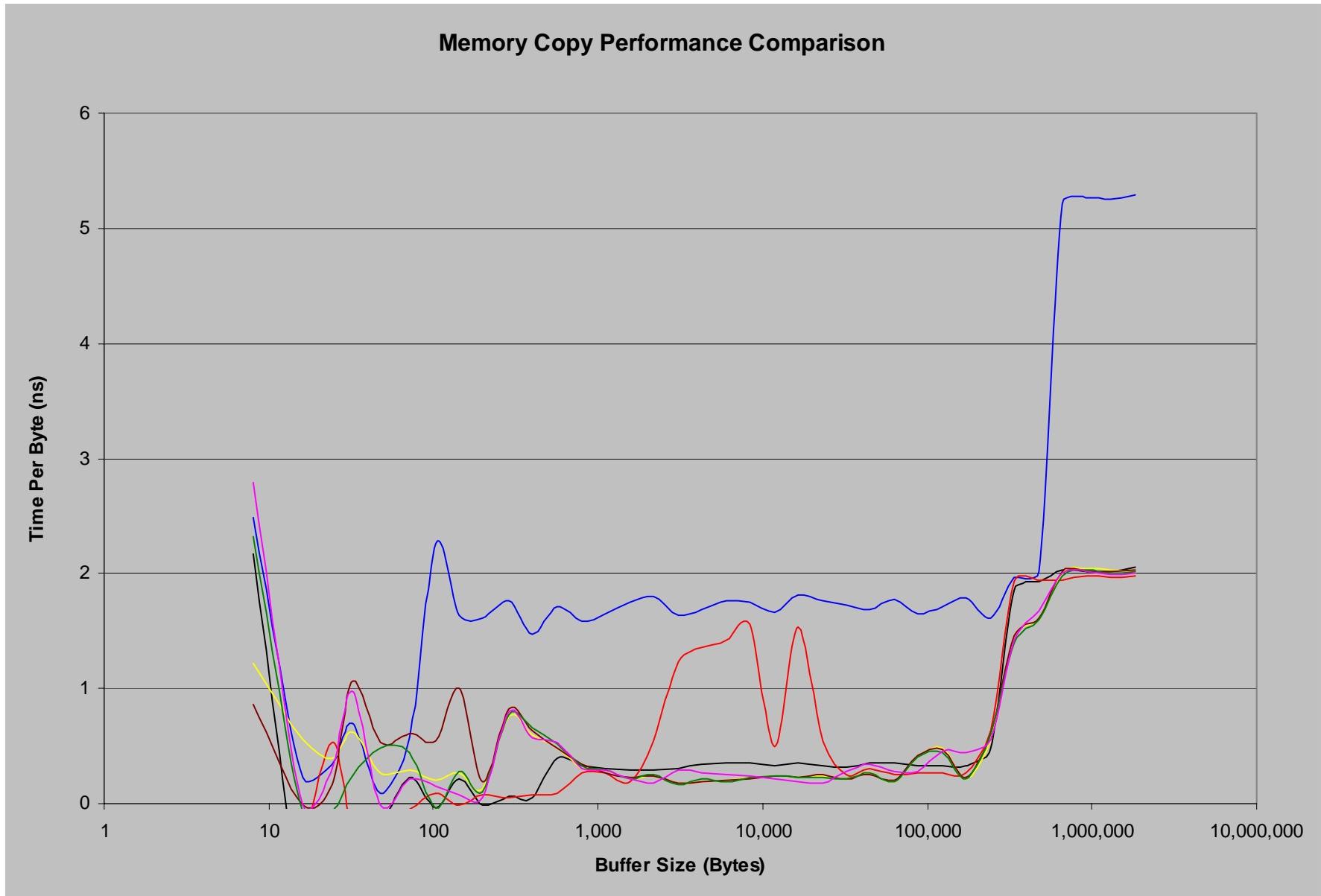
- ◆ **CPUs**
 - 2 GHz Pentium 4M laptop
 - 1 GHz Athlon desktop
 - (2 GHz P4M is 20% faster than 1 GHz Athlon)
- ◆ **Transports**
 - Shared memory on Windows
 - 100 Mb Ethernet



Memory Copy Performance

- ◆ Number of Algorithms
- ◆ Performance varies depending on:
 - Cache size
 - Cache line size
 - Bytes moved per operation
- ◆ ORB uses most efficient copy algorithm we can discover

Memory Copy Performance



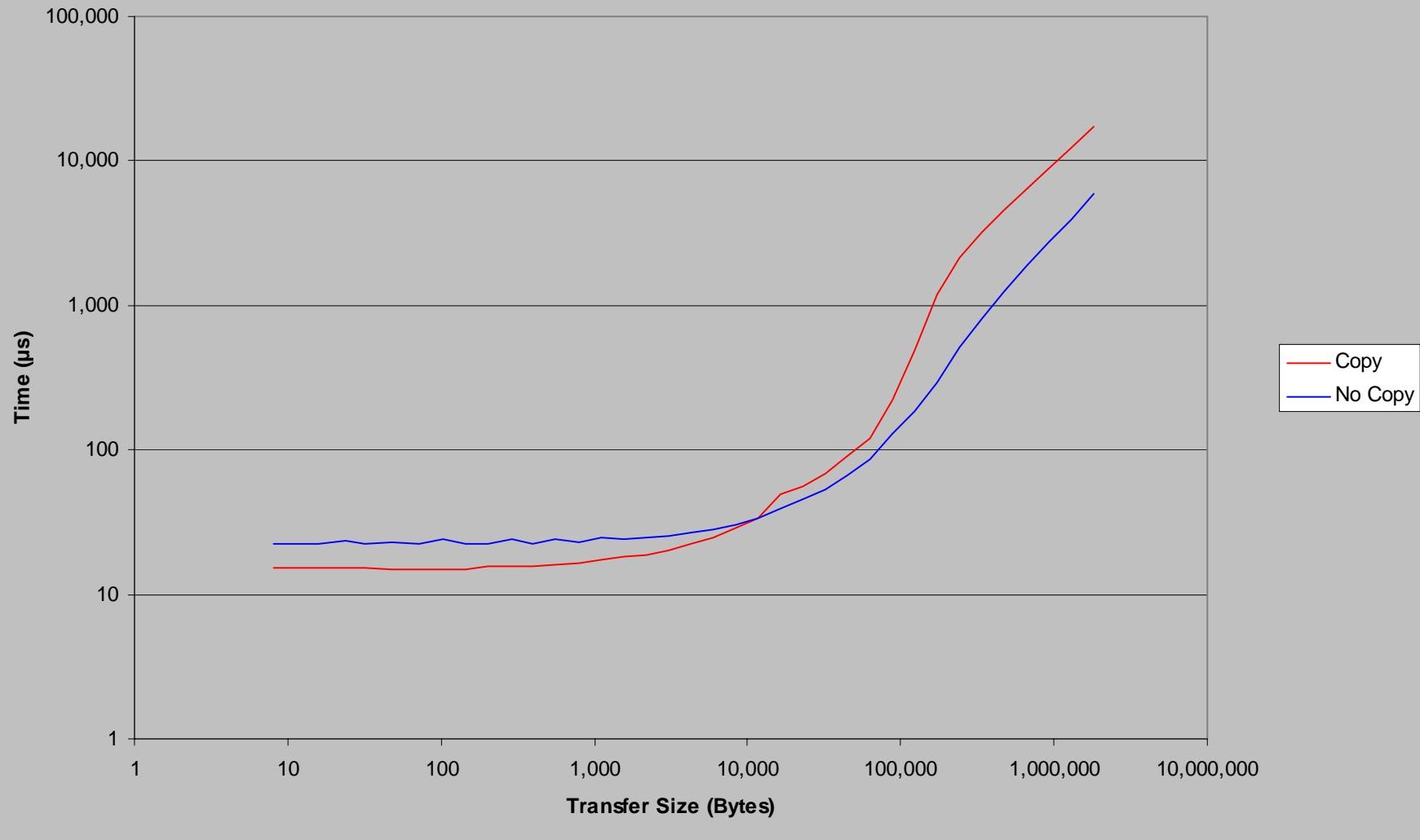


SHRMEM Latency

- ◆ Reducing marshaling copies
 - Decreases latency for large transfers
 - Increases latency for small transfers
- ◆ Latency increase occurs because there are more system calls from the transport
- ◆ Scatter/Gather system calls would reduce the number of calls, and potentially the number of transport copies.

SHRMEM Latency

Impact on Latency of Eliminating the Marshaling Copies
for the Shared Memory Transport



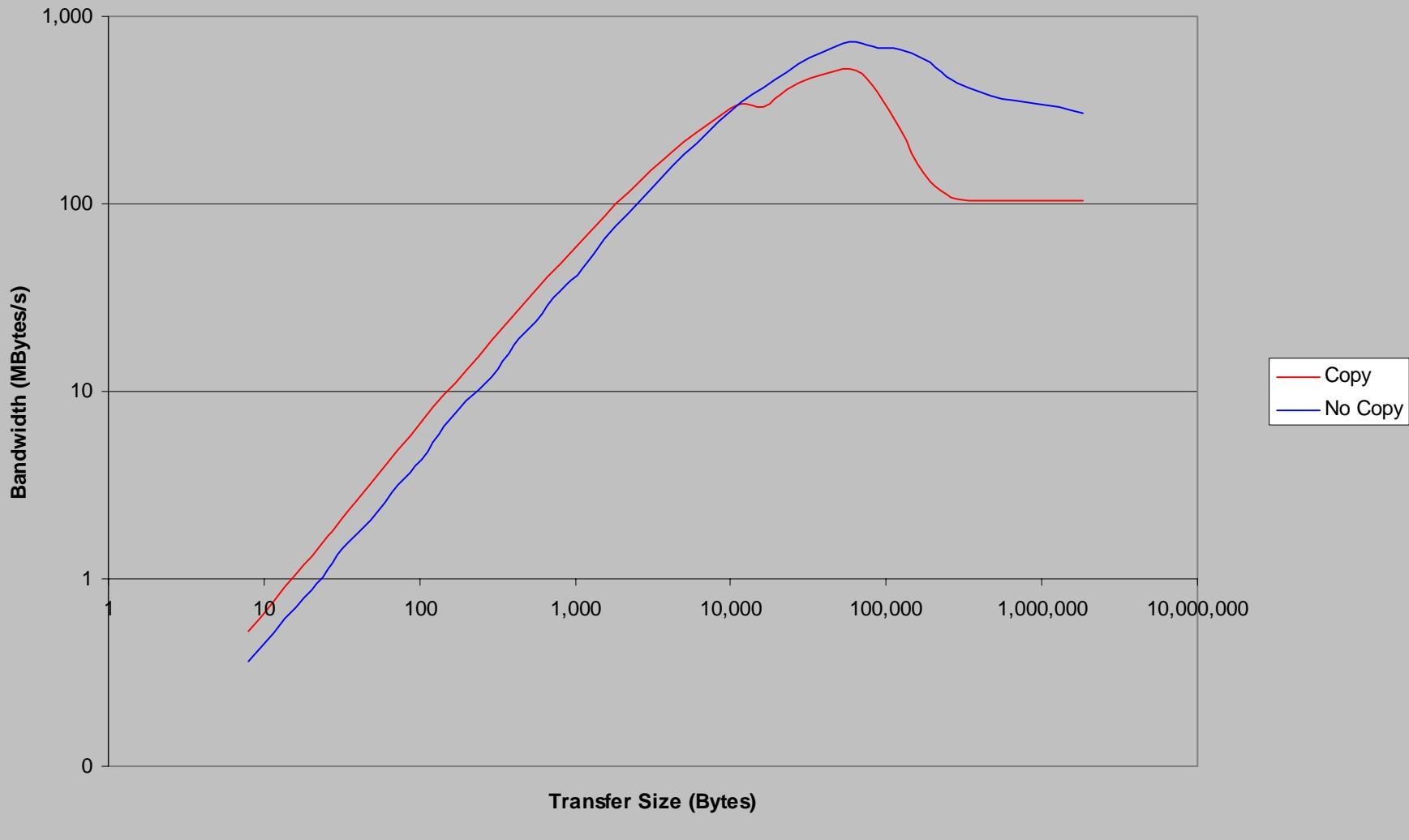


SHRMEM Bandwidth

- ◆ Reducing marshaling copies
 - Increases bandwidth for large transfers
 - Reduces bandwidth for small transfers
- ◆ Bandwidth reduction occurs because there are more system calls from the transport
- ◆ Scatter/Gather system calls would reduce the number of calls, and potentially the number of transport copies.

SHRMEM Bandwidth

**Impact on Bandwidth of Eliminating the Marshaling Copies
for the Shared Memory Transport**



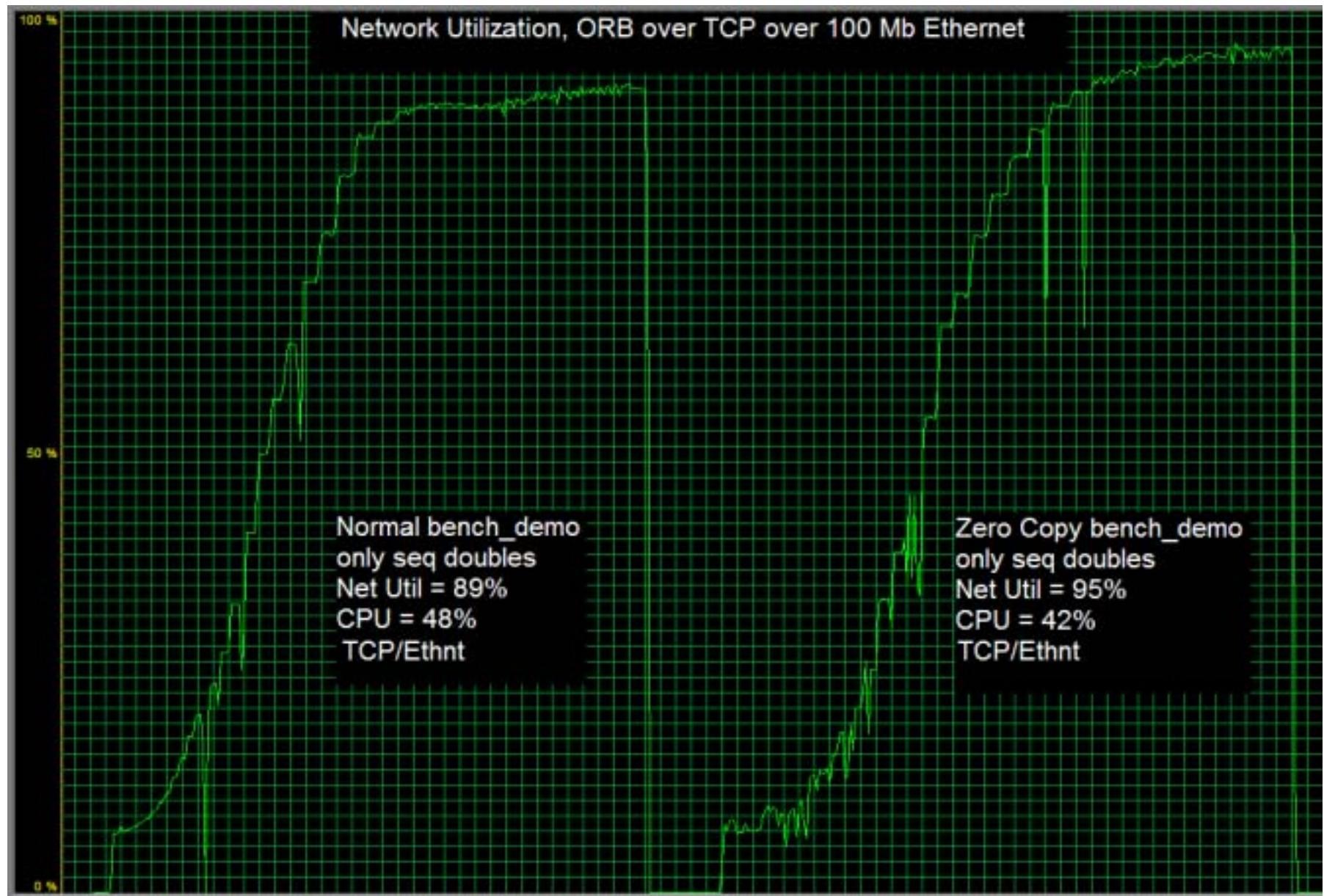
CPU Utilization

CPU Utilization (green: total CPU, red: kernel CPU)

non-ZC, over TCP/Ethnt, notice user CPU (green minus red) for larger data xfers

ZC, over TCP/Ethnt, virtually no user CPU used for larger data xfers

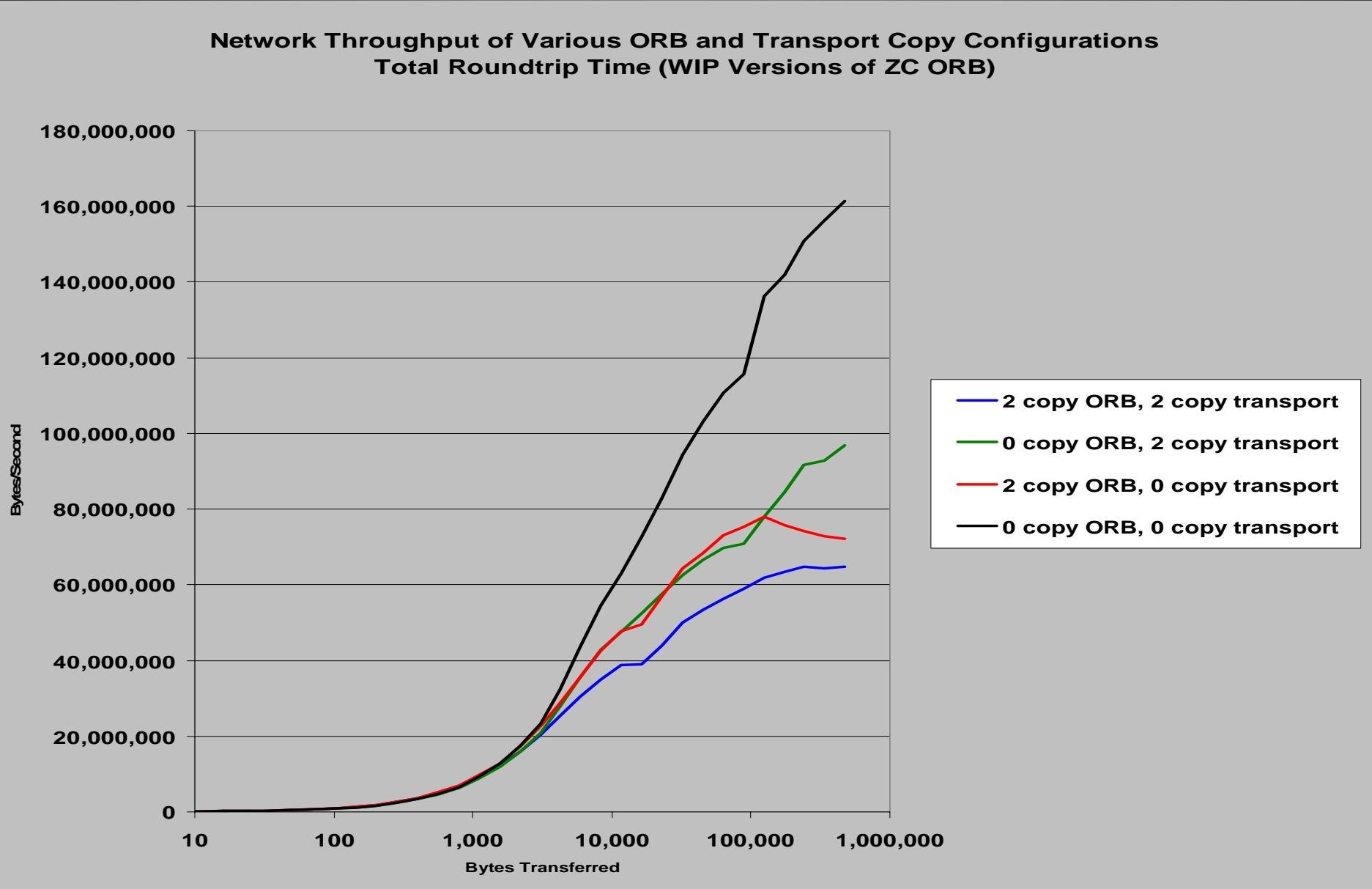
Network Utilization



First Benchmarks: Zero Copy Affect on HPEC

- ◆ Internal work-in-progress versions of ZC ORB
 - Several suboptimal characteristics
 - Underlying transport
 - ❖ High latency
 - ❖ DMA transfers are 80K blocks
- ◆ Mercury RACEway++
 - VxWorks host
 - CE-to-CE communications

Comparing Copy Configurations





Work Left

- ◆ Finalize Zero-Copy version of ORBexpress
- ◆ Rewrite underlying transport, expectations:
 - Better latency (> 10 usec)
 - More efficient use of DMA



Summary

- ◆ CORBA is progressing towards HPEC efficiency requirements
- ◆ Existing CORBA applications can take efficient advantage of HPEC hardware